



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

| APPLICATION NO.                   | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|-----------------------------------|-------------|----------------------|---------------------|------------------|
| 09/263,262                        | 03/05/1999  | MASAYUKI YAMADA      | 862.2720            | 1577             |
| 5514                              | 7590        | 10/04/2005           | EXAMINER            |                  |
| FITZPATRICK CELLA HARPER & SCINTO |             |                      | KNEPPER, DAVID D    |                  |
| 30 ROCKEFELLER PLAZA              |             |                      | ART UNIT            |                  |
| NEW YORK, NY 10112                |             |                      | PAPER NUMBER        |                  |

2654

DATE MAILED: 10/04/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

## Office Action Summary

Application No.

09/263,262

Applicant(s)

YAMADA, MASAYUKI

Examiner

David D. Knepper

Art Unit

2654

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1) ☒ Responsive to communication(s) filed on 25 July 2005.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4) ☒ Claim(s) 1-27 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-27 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
  - ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- |  |   |
|--|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892)   | 4) <input type="checkbox"/> Interview Summary (PTO-413)<br>Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)                                   | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152)             |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)<br>Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____  |

1. Applicant's correspondence filed on 25 July 2005 has been received and considered. Claims 1-27 are pending. The rejection below is drawn from the Final Office Action of 15 April 2003. Most of the applicant's arguments are towards arguments made by the Examiner in response to claim language that is no longer present and many quotes by the applicant are from arguments from the Office Action of 3 Jun 2004 which are repeated below for clarity.

### Specification

2. The change on page 1, lines 11-20 is objected to as new matter. The applicant's previous description admitted as "conventional" prior art the fact that it is well known to concatenate speech elements to include waveforms elements corresponding to "one to several pitches". The applicant is now trying to remove "one" which is not correct. It is well known to utilize a single pitch or multiple pitch values depending on a variety of factors such as the particular words or sentences (for example) that are being formed as well as the length of the particular elements that are being concatenated. Thus, it would be disingenuous for the applicant to rewrite history in such a way that something that has been available to the public might be construed otherwise given that the only criticism that applicant has for a waveform concatenation technique is the limitations imposed by pre-storing a "narrow range" of examples to be used to model prosody.

3. The disclosure is objected to because of the following informalities: In addition to run-on sentences, pages 2-8 contain claim language inappropriate for the body of the specification (see details below). Proper grammar and technical writing practices should be followed to make coherent statements in the specification. Appropriate correction is required.

The change on page 1, line 21 to page 2, line 1 is acceptable since it admits that it is well

known to search through speech data using more than one type of prior knowledge including the phonemic context of one or more phonemes or the pitch of the stored data. Thus, it is admitted that speech data is properly formed by selecting the concatenation one or more phonemes. It is also admitted that it is known to select such speech data based on pitch (fundamental frequency). Use of the selected data in combination will improve speech quality by employing these contextual and prosodic elements.

The change at page 2, line 22 to page 3, line 13 is confusing because: "The search means is the searching the database" is grammatically confusing; "The re-search." is not a complete sentence; and the last sentence has antecedent problems because of the confusing nature of the earlier sentences in the paragraph. To further prosecution, this material is presumed to represent admitted prior art.

The change at page 3, line 14 to page 4, line 14 does not match the marked-up copy. To further prosecution, this material is presumed to represent admitted prior art.

The change at page 4, line 21 to page 5, line 7 is objected to because "correspondence with these cond or third phoneme." does not make sense. To further prosecution, this material is presumed to represent admitted prior art.

The change at page 5, line 8 to page 6, line 8 still contains claims language and run-on sentences (in particular, the second sentence).

The change at page 6, line 9 to page 7, line 4 still contains claims language and run-on sentences (in particular, the second sentence).

The change at page 7, line 5 to page 8, line 8 still contains claims language and run-on sentences (in particular, the second and third sentences).

### Claims

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

5. Claims 1-27 are rejected under 35 U.S.C. § 103 as being unpatentable over Kaja (5,659,664) in view of Huang (5,913,193).

Claims 1, 12 and 23:

“Speech synthesis” is taught by Kaja’s speech synthesis, title:

“generating a second phoneme in consideration of a phonemic context for a first phoneme . . . searching said database for a phonemic piece data . . . re-searching said database for phonemic piece data corresponding to the third phoneme” (this is suggested by Kaja’s use of stored parameters used for triphone synthesis, which, by definition, relies on relationships between 3 phonemes (see Huang, col. 1, line 47). As stated in column 2, lines 63-65, Kaja relies upon the interconnection of several phonemes); and

“registering the search result . . . in a table” (suggested by his matrix, col.

2, line 16 – see Huang who explicitly teaches that his storage relies upon a table of senones stored in HMM storage 24, col. 3, lines 60-65).

It is noted that Kaja does not explicitly teach the use of a “table”. However, he teaches that his data must be stored for searching in a matrix. Huang explicitly teaches the use of a table as noted above. It would have been obvious for a person having ordinary skill in the pertinent art, at the time the invention was made, to use a table to store the result of searching the stored synthesis parameters because Huang teaches that determining the parameters for storage requires re-estimating the HMM parameters given the speech segmentation and that this will increase the probability of the HMM generating correct parameters. The desirability to generate correct parameters is the reason proper storage in a table is obvious.

Claims 2-27 are rejected under similar arguments as noted above. Utilizing pitch is taught by Kaja’s fundamental sound curve, col. 1, line 31 and Huang’s mean and variance for pitch, col. 7, line 27. Vowel and consonant combinations are obvious to anyone of ordinary skill in the art and are taught by Kaja in col. 2, lines 65 to col. 3, line 6 as well known triphone combinations. Any combination of 3 phonemes is anticipated by the term of art “triphone”.

Claims 8, 19 and 24: “The phoneme as a synthesis target”, “the acquired fundamental frequencies” and “phonemic context” is clearly taught by Huang’s stream of phonemes is transmitted to prosody engine 35, col. 8, line 19. Furthermore, Huang teaches in col. 8 that the context as determined for phonemes is affected by the proper intonation of a sentence, etc. to create natural sounding speech.

**Remarks from 15 April 2003 are in bold** (added remarks are in regular type)

6. Applicant's argument on page 19 that Kaja fails to disclose or suggest a "storage means for storing a table" is admittedly false in that the applicant acknowledges that Kaja does teach a matrix for storing polyphones. The applicant's argument that Kaja's polyphones are different than "storing triphones of the present invention" is contradicted by Kaja's preferred embodiment which uses triphone synthesis (col. 2, line 64).

The applicant's arguments of 25 July 2005, pages 14-16, that the cited passages of prior art are limited to col. 2, lines 63-65 (Kaja) and col. 1, line 47 (Huang) fails to address the additional arguments below from 3 Jun 2004 which points out that Kaja uses the terminology "polyphone" in his abstract as well as figures 1 and 2. Having been given these citations previously, the applicant should have addressed the terminology in Kaja's abstract which indicates: The behaviour of the respective parameter with time is defined around each phoneme boundary and polyphones are joined by forming a weighted mean value of the curves which are defined by their two associated matrices/sequences list. The applicant's focus on the "diphone" and "triphone" terms are misplaced because the applicant uses the broader terminology "polyphone". What Kaja teaches is that it is obvious to substitute diphones or triphones (narrow terms) for any interconnected polyphone (broader term) representations for synthesizing speech (see Kaja's figure 1, 2 and col. 2, lines 62-65).

The applicant's argument that Kaja fails to show the re-searching combination to rely "...in some way upon the context of adjacent phonemes" (applicant's response, page 15, 25 Jul 2005) was addressed with the arguments above in relation to storage. The applicant recognizes some relationship to Kaja (col. 3) on page 16 but fails to understand it properly. The storage is further described by Kaja in col. 3, lines 24-50 stating in part: ...an iterative process which, by

stages, ensures that the synthetic phrase more and more resembles the natural phrase. When a sufficiently good likeness has been obtained, the control parameters which correspond to the desired diphone/polyphone, can be extracted from the synthetic phrase. Further explanation of figure 1 by Kaja in column 3, lines 44-55 indicates that the combination of phonemes are done ...in time around the phoneme boundary (in FIG.1 and S3 in FIG. 2)... Two diphones are joined together (S4 in FIG. 2)... Thus, it should now be clear that the previous references to figures 1 and 2 of Kaja correspond to the claim limitations as further explained above because Kaja clearly teaches an iterative process (“re-search”) that is based upon phonemic context.

It is further noted that the reason the Examiner used the term “triphone” was because the applicant discloses this is the preferred embodiment for implementing the more broadly claimed “polyphones”. Thus, the Examiner was trying to use terminology that the applicant was familiar with (regarding well-known synthesis models that rely upon phonemic context) in order to show how closely related the prior art cited is, not only to the claimed invention, but to the disclosed invention. This was intended to help the applicant realize that much narrower claim terminology is needed to differentiate over the prior art of record.

**Applicant’s argument on page 19 that Huang fails to disclose the present invention which “manages a database for managing phonemic piece data and an index table having substitute phoneme data with respect to all the conceivable phonemic contexts” contradicts the teachings of Huang in column 3, lines 60-65 noted above which teach the use of a dictionary storage 22 which stores the phonemic description of each word and the HMM storage 24 which contains a table of senones. By definition, a senone is: An equivalence class which models a subphonetic event usually one state in a HMM for a phoneme**



(different phone models can share the same senone if they exhibit acoustic similarity). Therefore, it is clear that the claimed substitute phoneme data would read on the typical use of senones to allow different phone models to substitute for one another if they exhibit acoustic similarity. HMM's are particularly useful for modeling insertion, deletion and substitution. One of ordinary skill in the art would know that acoustic similarity of different phonemes occurs when they are in a similar context which is why modeling techniques such as described by Huang are used (see columns 5-6 of Huang).

The contextual modeling of Huang (column 8) manages phonemic data and the claims read upon it as combined with Kaja.

The applicant's arguments that Huang fails to show "phonemic context" (page 17 from 25 Jul 2005) are contradictory. The applicant admits that the senones of Huang represent Markov states across phonetic models. This clearly teaches that they represent phonemic context. Thus, the interconnected elements taught by Kaja could easily be implemented using Markov models because both are utilized to improve the representation of speech by modeling phonetic information in context rather than other representations that are not sophisticated enough to represent transitions.

Regarding arguments on page 20 (25 Jul 2005), see Huang, col. 8, lines 38-39 which explicitly says: The pitch of a phoneme can be affected by the intonation of the sentence. Thus, the previous reference to col. 8 was an accurate representation of Huang's teaching.

The Examiner remarks from 3 June 2004 are repeated below to show a consistent effort to help the applicant understand the relationship between well-known terms of art.

**Remarks (from 3 June 2004)**

7. The changes to the claims make them broader than they were previously and therefore do nothing to overcome the prior art upon which the narrower claims presented earlier read.

As earlier presented, the definition of “phoneme” is **a member of the set of the smallest unites of speech that serve to distinguish one utterance from another in a language or dialect.**

The term now used is “phonemic label”. The definition of “phonemic” is **of, relating to, or having the characteristics of a phoneme.** The term “label” is defined as **to describe or designate with a label.** Therefore, the terminology has been modified by the applicant to broaden the scope of the claims to include certain speech sounds that have characteristics of phonemes regardless of how they are specifically modeled so long as they may be described as having some relationship to a phoneme.

The current claim language has substituted, “second phoneme” with --first label--, “first phoneme” with --phonemic label--; and “third phoneme” with --second label--. The previous rejection is not overcome because Kaja teaches control parameters are stored in a matrix or a sequence list for each polyphone (abstract, figures 1 and 2) which teaches the use of labeled parameters which represent phonemic data in the form of polyphones. As previously explained, Huang teaches a specific labeling technique using Hidden Markov Models (Fig. 3B) stored in a table (col. 3, lines 60-65). Thus, the prior art clearly teaches that it was well known in the art to synthesize speech by joining labeled phonemic data that was stored in a table.

The applicant’s arguments that the prior art fails to utilize contextual modeling is without merit. Diphones and triphones models inherently rely upon the context of adjacent phonemes.

Thus, the applicant's arguments display a lack of consideration or understanding of basic terms of art that one of ordinary skill in the art of speech synthesis would be intimately familiar with.

8. All claims are drawn to the same invention claimed in the application prior to the entry of the submission under 37 CFR 1.114 and could have been finally rejected on the grounds and art of record in the next Office action if they had been entered in the application prior to entry under 37 CFR 1.114. Accordingly, **THIS ACTION IS MADE FINAL** even though it is a first action after the filing of a request for continued examination and the submission under 37 CFR 1.114. See MPEP § 706.07(b). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire **THREE MONTHS** from the mailing date of this action. In the event a first reply is filed within **TWO MONTHS** of the mailing date of this final action and the advisory action is not mailed until after the end of the **THREE-MONTH** shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than **SIX MONTHS** from the mailing date of this final action.

9. Some correspondence may be submitted electronically. See the Office's Internet Web site <http://www.uspto.gov> for additional information.

Please address mail to be delivered by the United States Postal Service (USPS) as follows:

Mail Stop \_\_\_\_\_  
Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

**The Central fax number is 571-273-8300.** Please label INFORMAL" or "DRAFT" communications accordingly.

Mail Stop should be omitted if none is indicated.

Effective 14 January 2005, except correspondence for Maintenance Fees, Deposit Accounts (see 37 CFR 1.25(c)(4)), and Licensing and Review (see 37 CFR 5.1(c) and 5.2(c)), please address correspondence delivered by other delivery services (i.e. – Federal Express (Fed Ex), UPS, DHL, Laser, Action, Purolator, etc.) as follows:

U.S. Patent and Trademark Office  
Customer Window, Mail Stop \_\_\_\_\_  
Randolph Building  
Alexandria, VA 22314

10. Any inquiry concerning this communication or earlier communications from the examiner should be directed to David D. Knepper whose telephone number is (571) 272-7607. The examiner can normally be reached on Monday-Thursday from 07:30 a.m.-6:00 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil, can be reached on (571) 272-7602.

For the Group 2600 receptionist or customer service call (571) 272-2600.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Inquiries regarding the status of submissions relating to an application or questions on the Private PAIR system should be directed to the Electronic Business Center (EBC) at 866-217-9197 (toll-free) or 703-305-3028 between the hours of 6 a.m. and midnight Monday through Friday EST, or by email at [ebc@uspto.gov](mailto:ebc@uspto.gov). For general information about the PAIR system, see <http://pair-direct.uspto.gov>.



David D. Knepper  
Primary Examiner  
**Art Unit 2654**  
September 27, 2005